



Deep learning provides high accuracy in automated chondrocyte viability assessment in articular cartilage using nonlinear optical microscopy

XUN CHEN,^{1,2} YANG LI,¹  NICOLE WYMAN,¹ ZHENG ZHANG,¹ HONGMING FAN,¹ MICHAEL LE,¹ STEVEN GANNON,¹ CHELSEA ROSE,¹ ZHAO ZHANG,¹ JEREMY MERCURI,¹ HAI YAO,¹ BRUCE GAO,¹ SHANE WOOLF,³ THIERRY PÉCOT,⁴ AND TONG YE^{1,5,*} 

¹Department of Bioengineering, Clemson University, Clemson, SC 29634, USA

²Current address: Institute of Medical Photonics, Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing 100083, China

³Department of Orthopedic, Medical University of South Carolina, Charleston, SC 29425, USA

⁴Hollings Cancer Center, Medical University of South Carolina, Charleston, SC 29425, USA

⁵Department of Regenerative Medicine and Cell Biology, Medical University of South Carolina, Charleston, SC 29425, USA

*ye7@clemson.edu

Abstract: Chondrocyte viability is a crucial factor in evaluating cartilage health. Most cell viability assays rely on dyes and are not applicable for *in vivo* or longitudinal studies. We previously demonstrated that two-photon excited autofluorescence and second harmonic generation microscopy provided high-resolution images of cells and collagen structure; those images allowed us to distinguish live from dead chondrocytes by visual assessment or by the normalized autofluorescence ratio. However, both methods require human involvement and have low throughputs. Methods for automated cell-based image processing can improve throughput. Conventional image processing algorithms do not perform well on autofluorescence images acquired by nonlinear microscopes due to low image contrast. In this study, we compared conventional, machine learning, and deep learning methods in chondrocyte segmentation and classification. We demonstrated that deep learning significantly improved the outcome of the chondrocyte segmentation and classification. With appropriate training, the deep learning method can achieve 90% accuracy in chondrocyte viability measurement. The significance of this work is that automated imaging analysis is possible and should not become a major hurdle for the use of nonlinear optical imaging methods in biological or clinical studies.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Chondrocyte viability is a crucial factor in evaluating cartilage health. Prevalent cell viability assays rely on dyes and are not applicable for *in vivo* or longitudinal studies [1,2]. Recently, we demonstrated that two-photon excitation autofluorescence (TPAF) and second harmonic generation (SHG) microscopy provided high-resolution images that can distinguish live/dead chondrocytes in the articular cartilage tissue [3]. The majority of TPAF in cells originates from the reduced form of nicotinamide adenine dinucleotide (NAD) or nicotinamide adenine dinucleotide phosphate (NADP) and flavin proteins (FPs); collagen fibrils yield both TPAF and SHG signals in the extracellular (ECM) region. TPAF and SHG are both intrinsic signals from endogenous molecules that are available in cartilage tissue. As such, our TPAF/SHG chondrocyte viability assay [3] does not need to introduce any labeling dyes to samples, allowing for the

assessment of cartilage tissue in a non-contact fashion and possibly *in vivo* if an appropriate imaging device is developed. In this nonlabeling assay, the cell status is classified by either the visual observation of the multichannel, pseudo-color images or the cell-based quantitative analysis with the normalized autofluorescence ratio [3] upon manual cell segmentation. Both methods rely on human involvement and their throughputs are low. Methods for automated cell-based image processing are necessary to improve the throughput of chondrocyte viability assessment for cartilage studies.

Chondrocyte viability is defined as the percentage of live cells in the total cell population. Automated viability assessment needs to determine both populations for the calculation. In general, three major imaging processing tasks, including segmentation, detection and classification, are involved in the method. Segmentation and detection separate cellular areas from the ECM area and identify individual cells; classification determines if a cell is alive or not. Segmentation, detection and classification are common imaging processing tasks in the cell-based image analysis. Many algorithms have been developed to complete these tasks. Readers can refer to the listed Refs. [5] for reviews of these algorithms and their uses in cell-based image processing. Recent advances in deep learning (DL) algorithms have significantly leveraged the competency of automated cell-based image processing in the microscopy field [6,7]. Both the accuracy of analysis and the complexity of tasks have significantly increased in comparison with what conventional, non-deep-learning algorithms can provide. One of the major advantages of deep-learning-based image processing is that features or patterns used in segmentation and classification are not pre-defined; instead, a comprehensive training process is required to establish networks to process images using a large number of images acquired under similar settings. In contrast, conventional algorithms do not need the training process, but pre-defined features are essential. For example, in conventional cell segmentation [4] and detection methods [5], the pixel intensity and its distribution patterns often serve as thresholds or morphological features to identify cellular areas. In cell classification, quantitative measures must be defined as criteria to determine the category (e.g., live vs dead, or cancerous vs non-cancerous) of a cell. Although conventional methods are easier to implement and more efficient in the utility of computing resources, their accuracy is often low, reflected in the cell-touching problem (being not able to isolate individual cells) in segmentation and by inaccurate cell counts in classification. Automated chondrocyte viability assessment is a challenging task; perfect cell segmentation is also difficult with the conventional algorithms due to the low image contrast of TPAF images and densely packed chondrocytes in the superficial zone. However, we hypothesize that the deep learning algorithms may provide higher accuracy than conventional methods in automated chondrocyte viability analysis.

DL algorithms have been successfully demonstrated in fields such as biological and medical image processing [8,9]. In the cell-based analysis, a few studies used DL in either segmentation [9] or classification [10]. Yang et al. demonstrated a DL method used for automated chondrocyte identification on histological slides of articular cartilage [11]. It was demonstrated that U-Net, one of the DL networks, could achieve superior performance compared to conventional methods in cell nuclei segmentation [12]. The U-Net network was able to achieve an accuracy score between 0.6 and 0.8 in the general pixel-based classification for cell counting and detection; a prediction and ground truth intersection overlay also reached around 0.9 in morphometry segmentation [8]. Yu et al. compared another DL network, AlexNet, with conventional machine learning (ML) methods in the multivariate classification of liver fibrosis stages [13]. It was demonstrated that DL was significantly better than conventional ML in assessing texture features. To our knowledge, DL networks have not been demonstrated in the chondrocyte viability assessment, a task that involves both segmentation and classification.

In this report, we present a deep learning strategy for the chondrocyte viability analysis. The U-Net model was demonstrated to improve the quality of segmentation. However, due to the cell touching problem, a perfect segmentation was still hard to achieve, which caused significant

errors in cell counting and led to inaccuracy in the viability analysis. To solve the problem, we proposed a dual multivariate classification using CNN to evaluate the live and total cell populations separately in two independent classifications. This method does not require perfect detachment of each individual cell. We have demonstrated that the deep learning multivariate cell classification provides over 90% accuracy in viability analysis, which is significantly better than other methods such as machine learning and conventional algorithms. Although our methods are particularly developed for chondrocyte viability assessment, the strategy of deep-learning-based imaging processing may also be used in many other similar applications.

2. Materials and methods

2.1. Image acquisition

Tibias from male Sprague-Dawley rats ($n=15$) were acquired according to guidelines of the Institutional Animal Care and Use Committee. Cartilage samples were prepared and imaged according to the protocols described previously [3]. The imaging microscope was a commercial multiphoton optical microscope (FV1200 inverted, Olympus Corporation, Tokyo, Japan) equipped with a Ti:Sapphire ultrafast laser (MaiTai Deepsee, Newport, CA). Two GaAsP photomultiplier tubes (PMTs) were used to acquire two imaging channels simultaneously. TPAF images were taken when the laser was tuned to 740 nm. Simultaneously, autofluorescence of NAD(P)H and FPs was collected by the violet (420–460 nm) and red (575–630 nm) channels, respectively. The voltage of both PMTs was set at 650 V while the digital gain and the offset of preamps were set at 1.0 and 0, respectively. A 30 \times silicone immersion objective lens (UPLSAPO 30x, Olympus) with 1.05 N.A. was used in the imaging experiments. The size of the acquired images was 1024 \times 1024 pixels, corresponding to a field of view of 423 $\mu\text{m} \times$ 423 μm . The pixel dwell time was set at 2 μs with the Kalman filter online integration of 2. Three-dimensional images were acquired in the z-stack mode with 50 slices/stack and 1 $\mu\text{m}/\text{step}$; the thickness of the stack covered roughly the entire superficial layer of the cartilage tissue. SHG images were taken right after the TPAF imaging with the laser tuned to 860 nm. The violet channel was used to collect the SHG signal at 430 nm. The TPAF and SHG stacks were merged to form a three-channel stack for each sample region.

2.2. Ground truth definition for segmentation and classification

For the visual assessment of cell viability, using ImageJ (Fiji) [14], a single RGB-color image was formed by merging three channels of an acquired image with red assigned to the red fluorescence channel (FPs), green to the violet fluorescence channel (NAD(P)H) and blue to the SHG channel. Figure 1 shows a typical acquisition of TPAF and SHG channels and the corresponding RGB color image. The viability of each cell in an image was assessed visually by evaluating the brightness and color appearance. Bright and green cells were live while dim and red cells were dead. The ground truth for cell segmentation was obtained by manually tracing the edge of each lacune in the SHG images using ImageJ (Fiji) on a graphic drawing tablet (Wacom Intuos Pro). An image with viable labels were first segmented by using different automated segmentation methods and then divided into classes by visual judgment. For each image, classification results were stored in a database used to determine the ground truth for viability. We measured 15 tissues in articular cartilage. The training/testing sets came from the same tissue, and the viability validation set for reported results herein came from different tissues.

2.3. Detection and segmentation of chondrocytes

Generally, TPAF channels show less contrast than the SHG channel. Consequently, SHG images are used for detection and segmentation of the chondrocyte region. In cartilage, chondrocytes are surrounded by a thin layer of the pericellular matrix (PCM), which separates cells from ECM.

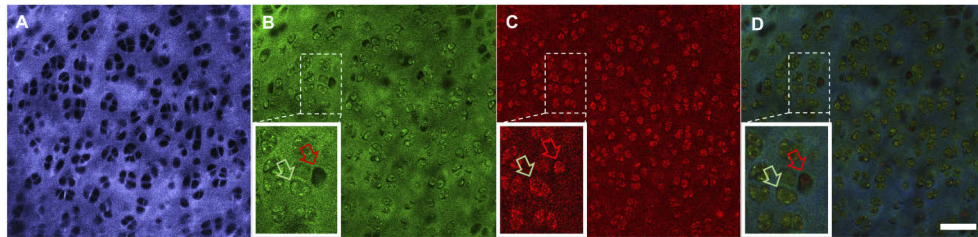


Fig. 1. Typical images acquired from cartilage tissue by a nonlinear microscope: (A) SHG; (B) Violet fluorescence; (C) Red fluorescence; and (D) Merged RGB (Red: red fluorescence; Green: violet fluorescence; Blue: SHG) image used for chondrocyte viability assessment. Green arrows point to live cells and red point to dead cells. Scale bar: 50 μm .

Fibrillar collagens (mainly type II) are located only in the ECM area and are the primary source of SHG signals in cartilage. Chondrocytes and the surrounding PCM appear as dark holes (lacunae) in the SHG channel. This typical appearance of chondrocytes suggests that SHG images can be used to segment the lacuna area to form binary masks that can extract cellular areas in TPAF channels for classification. To generate those masks, SHG images were converted to inverted grayscale, and the lacuna areas were then segmented using different algorithms. Outcomes of the segmentation were binary masks to extract cellular areas in TPAF images. Figure 2(A) summarizes the workflow briefly.

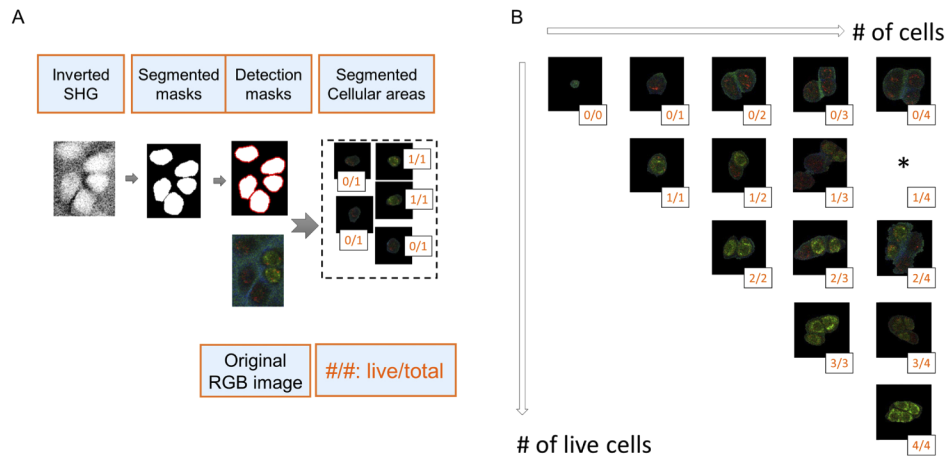


Fig. 2. Chondrocyte viability assessment workflow. (A) Cell segmentation, detection, and classification. (B) Cell multivariate classification with five outcome categories (0, 1, 2, 3, and 4) according to the number of total cells or number of live cells contained in a segmented cellular area. *: no image was found in our current dataset.

Cell segmentation is a common image processing task in fluorescence microscopy, and many segmentation methods, such as thresholding, morphology and watershed [15], have been developed. DL-based approaches have recently demonstrated a better performance than conventional methods [16]. The ability to learn image features that are suited to a given task makes these approaches particularly efficient. To identify the best approach for chondrocyte segmentation, we compared two conventional methods, adaptive thresholding/blob detection (ATB) [17] and watershed segmentation (WS), to the DL method U-Net (UN) [8] and a hybrid algorithm that applies the watershed method to the U-Net results (UW) [18]. ATB and WS were implemented using an Open Source Computer Vision Library (*OpenCV 4.2.0*). The U-Net model

(See Supplementary) was implemented with *Keras 2.1.6* library and *Tensorflow 1.4.0*. UN was trained by using over 4000 inverted SHG images (256×256 pixels) with manually annotated cellular areas. The training procedure took about 20 minutes on a customized computer equipped with a single graphics card (EVGA GeForce GTX 1080Ti FTW3) and an Intel i9-7920x CPU. [Code 1](#) written in Python is hosted at GitHub [19].

2.4. Dual multivariate classification for the chondrocyte viability assessment

The cellular areas obtained from SHG images are used in the RGB images to evaluate the live and dead cells. The chondrocyte viability is defined as the percentage of live cells over the total number of cells. With a perfect cell segmentation, viability is easy to compute after cell classification, for which only two categories, live and dead, are needed. Unfortunately, segmented cellular areas often contain more than one cell, which creates a significant error in the viability computation. To address this problem caused by inaccurate cell segmentation, we propose a dual-multivariate-classification strategy. Two independent classifications are performed, and each classification uses five categories instead of two (live and dead). In one classification, five categories are defined by the total number of cells (0, 1, 2, 3, and 4) contained in each segmented cellular area. In the other classification, five other categories are defined by the number of live cells (0, 1, 2, 3, and 4) contained in each segmented cellular area. Five categories are used because the probability of more than 4 chondrocytes being found in a single segmented cellular area is extremely low (none has been found so far). If over 4 cells were found, the program would count 4 cells. Once the number of live and total cells are determined by the classification processes, the chondrocyte viability of an image can be determined by using Eq. (1). Figure 2 represents the workflow of the proposed automated chondrocyte assessment, including the segmentation procedure and the dual-classification method with two sets of cell categories.

$$Viability = \frac{\sum_{i=1}^N N^{(i)}_{Live}}{\sum_{i=1}^N N^{(i)}_{Total}} \quad (1)$$

where i is the index of segmented cellular areas included in a viability analysis, with a maximum of N ; N_{Live} and N_{Total} are the number of live cells and the total number of cells, respectively, in the i th cellular area.

Cell classification was performed with the k-nearest neighbors (KNN) algorithm [20], a support vector machine (SVM) [21] model, and a convolutional neural network (CNN) [22]. KNN and SVM are conventional machine learning methods that use histogram orientation gradient (HOG) [23], morphology detected circles, pseudo color histogram [24] and spatially binned features as input image features for classification. They were implemented with the standard machine learning library *SKlearn 0.21* [25]. The CNN architecture used for cell classification was an AlexNet containing 3 convolutional layers and 2 fully connected layers as shown in Fig. S2(B). The AlexNet was implemented with the *Keras 2.1.6* library and *Tensorflow 1.4.0*. As shown in Fig. 2(B), 15 categories were defined for cell classification if combination of live and total cell number are considered. A single AlexNet with 15-category classification should work. However, we trained two independent AlexNet networks, each of which performed a classification with five categories (0, 1, 2, 3 and 4), one for the live cell numbers, and the other for the total cell number on input images. The two AlexNet networks were trained and validated separately without any dependency to each other. This dual 5-category classification strategy reduced the training time and improved the accuracy of classification. 6000 images were included in the training/testing data set. The training procedure took about 30 minutes on a customized computer equipped with a single graphic card (EVGA GeForce GTX 1080Ti FTW3) and an Intel i9-7920x CPU. [Code 1](#) is written in Python and is currently hosted at GitHub. [19].

2.5. Evaluation measures

F1 score was used to evaluate the performance of chondrocyte segmentation/detection:

$$F1 = 2 \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}, \quad (2)$$

where

$$\text{precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}}, \quad (3)$$

$$\text{recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}}. \quad (4)$$

In the above equations, true positives correspond to correct cellular area segmentation/detection compared with the ground truth of visual judgment. False positives correspond to false cellular area segmentation/detection. False negatives correspond to false non-cellular area segmentation/detection. Outcomes of cell classifications were evaluated with respect to sensitivity and specificity as follows:

$$\text{Sensitivity} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}}, \quad (5)$$

$$\text{Specificity} = \frac{\text{True negative}}{\text{True negative} + \text{False positive}}, \quad (6)$$

where true positive corresponds to the number of correct live cells compared with the ground truth according to the visual judgment, true negative corresponds to the number of correct dead cells, and false positive correspond to the number of false live cells.

The multivariate classification for live/total chondrocyte numbers was evaluated with a multi-class receiver operating characteristic (ROC) analysis [26] according to the method described in this website [27]. By using trained AlexNet, probabilities of cell number categories (0, 1, 2, 3, and 4) were predicted. A ROC curve was generated by continuously varying the threshold of the probability for each category based on the ground truth. The area under the ROC curve (AUC) ranging from 0 to 1 evaluates the ability of a model to accurately distinguish different categories.

After the total number of cells and the number of live cells were defined, a viability ratio ranging from 0 to 1 was calculated for each TPAF/SHG image. The mean squared error (MSE) was used to evaluate the overall outcome of the viability assessment and is defined by:

$$MSE = \frac{1}{n} \sum_{i=1}^n \sqrt{(R_G - R_P)^2}, \quad (7)$$

where R_G is the ground truth as manually defined, and R_P is the model prediction for the n cell clusters. The root squared error (RSE) for each TPAF/SHG image for the evaluation in Fig. 6 was computed as follows:

$$RSE_i = \sqrt{(R_{Gi} - R_{Pi})^2}, \quad (8)$$

where R_{Gi} and R_{Pi} are the ground truth and predicted viability ratios for the i th image, respectively. Descriptive statistics were reported as mean standard deviation. Statistically significant differences were reported when $p < 0.05$. Statistical analysis was performed using GraphPad (GraphPad Software, Inc).

3. Results

3.1. Segmentation of cellular areas on SHG images

The first step for automatically assessing chondrocyte viability is to segment individual cells. As described in Methods, SHG images provide higher contrast than autofluorescence channels and are, therefore, better suited to identify cellular areas (lacunae). In this study, we compare the performance of four algorithms that are commonly used for cell segmentation. The first two algorithms are conventional methods, adaptive thresholding with blob detection (ATB) [17] and watershed segmentation (WS). The third approach is the convolutional neural network U-Net (UN) [8], the most popular deep learning method in biomedical imaging. The last method is a U-Net combined with watershed (UW) [18]. Figure 3(A) shows 3 examples of inverted SHG images and their segmentation results using the different methods. A sum of 30 images of the size of 256×256 pixels, selected from 30 different image stacks, are used for evaluation. The F1 scores are shown in Fig. 3(B). UW obtains the best score (0.92 ± 0.02), while UN, WS and ATB methods obtain F1 scores equal to 0.78 ± 0.03 , 0.77 ± 0.04 and 0.52 ± 0.03 , respectively. These results demonstrate that UW significantly improves cell segmentation in comparison to the three other approaches.

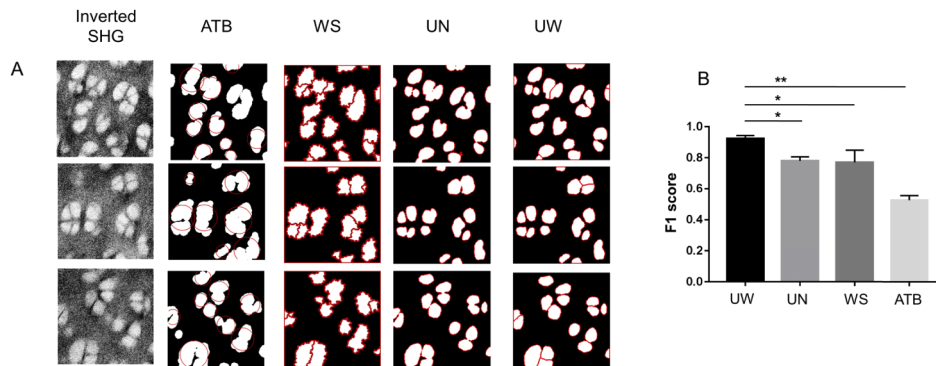


Fig. 3. (A) Three typical inverted SHG images and their segmentation results using four different algorithms. ATB: adaptive thresholding with blob detection; WS: watershed segmentation; UN: U-Net segmentation; and UW: U-Net segmentation followed by watershed segmentation (B) Comparison of the performance using F1 scores (*: p-value<0.05; **: p-value<0.01).

While UW obtains high accuracy segmentations of chondrocytes, the algorithm cannot completely separate each individual cell as is reflected in its F1 score (0.92 ± 0.02). For instance, areas #9, #14 and #15, evaluated with UW in Fig. 4, contain 2 cells each. Even though more cells are identified with UW than with the other methods, this failure to accurately identify each individual cell can potentially significantly alter the performance of chondrocyte viability assessment. Instead of trying to improve cell segmentation, we propose to use a multivariate classification strategy to estimate the correct number of cells from the segmented cellular areas.

3.2. Dual multivariate classifications of chondrocytes using conventional machine learning and deep learning methods

To correct for segmentation errors where cells are not separated accurately, we propose a dual multivariate classification strategy. We utilized the K-Nearest Neighbor method (KNN) [20], a Support Vector Machine (SVM) model [21] and a Convolutional Neural Network (CNN) [22] for the estimation of the total number of cells and the estimation of total live cells on a set of images (n=100) used for the segmentation evaluation. For each image, the 3 classification algorithms

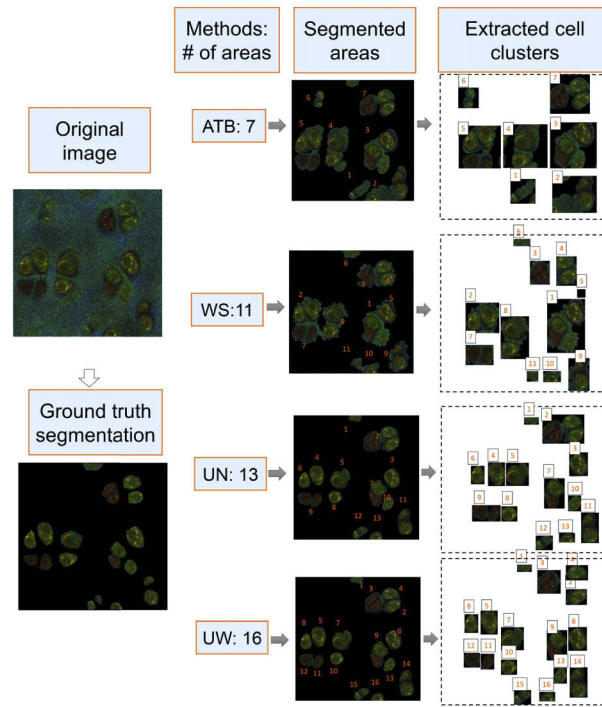


Fig. 4. Segmented chondrocytes from an SHG image using ATB, WS, UN, and UW. The number of chondrocytes as manually defined is 20. ATB, WS, UN and UW identified 7, 11, 13, and 16 chondrocytes, respectively.

were applied to the 4 segmentation results obtained with the 4 different segmentation algorithms, resulting in 12 classification outcomes. Multivariate ROC analysis [26] was performed for each classification. Figure 5 shows the ROC curves when plotting the true positive rate (TPR, i.e., Sensitivity) against the false positive rate (FPR, i.e., 1-Specificity) at various threshold settings. The best sensitivity, specificity, and area under the ROC curve (AUC) of multivariable classifications were calculated and are summarized in Table 1. Note that Table 1 lists only the best sensitivity, specificity, and AUC values among the four results yielded from 4 segmentation algorithms. Again, deep learning outperforms other conventional machine learning algorithms in both the live cell and total cell classification. The AUCs (1.0 is the best) of both classifications are over 0.9. These results demonstrate that the proposed deep learning classification method provided a promising strategy to overcome the problem caused by imperfect segmentation.

Table 1. The sensitivity, specificity, and AUC of live/total multivariate classifications.

Classification	Methods	Sensitivity	Specificity	AUC
Live cells	KNN	0.63 ± 0.04	0.91 ± 0.01	0.77 ± 0.03
	SVM	0.65 ± 0.04	0.90 ± 0.02	0.76 ± 0.04
	CNN	0.90 ± 0.03	0.93 ± 0.01	0.92 ± 0.02
Total cells	KNN	0.60 ± 0.05	0.91 ± 0.01	0.75 ± 0.03
	SVM	0.62 ± 0.03	0.92 ± 0.02	0.75 ± 0.02
	CNN	0.91 ± 0.04	0.92 ± 0.02	0.94 ± 0.01

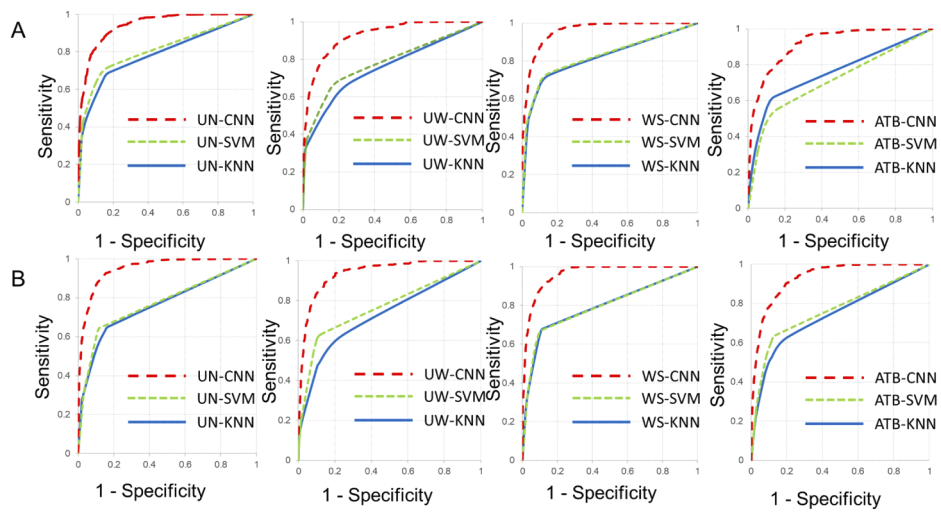


Fig. 5. The ROC curves of chondrocyte live/total cell number classification using CNN, KNN and SVM from four segmentation methods (UN, UW, WS and ATB). Sensitivity and specificity of multivariate classifications of live/total cell numbers. (A) Live cell number classification. (B) Total cell number classification.

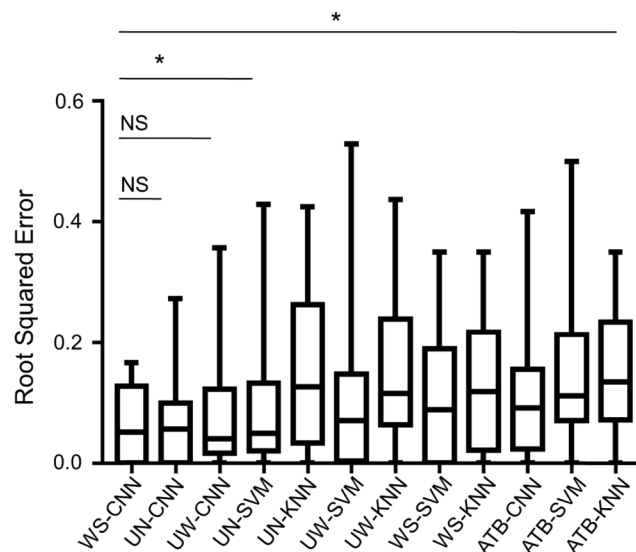


Fig. 6. Root squared errors of viability analysis using Deep/machine learning cell number ratio-based methods, combining chondrocyte detection methods ATB, WS, UN, and UW, and classification methods CNN, KNN and SVM. Statistically significant differences were reported at $p < 0.05$. (ns: no significant difference; *: p -value < 0.05 ; **: p -value < 0.01)

3.3. Comparing the accuracy of viability ratios determined by different segmentation-classification methods

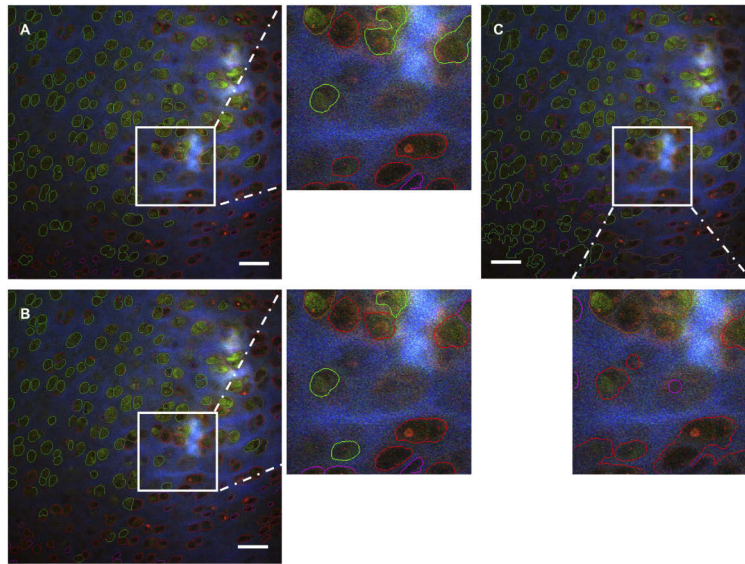


Fig. 7. Results of viability analysis of experimental images (1024×1024 pixels). (A) UN-CNN (B) UW-CNN (C) WS-CNN. Red circles: dead cells; green circles: live cells; yellow circles: both live and dead cells; magenta circles: non-cellular structures. Scale bar: 50 μm .

When the total number of cells and number of live cells were defined, the chondrocyte viability ratio was computed as described in Eq. (1). Then, the mean squared errors between the ground truth and model predictions were calculated using Eq. (7). To evaluate the overall performance of various algorithms in the viability analysis, we included 30 raw RGB images (1024×1024 pixels) acquired from TPAF/SHG label-free imaging experiments in the testing set. In this analysis, the raw images were sectioned into 16 sub-images (256×256 pixels) for segmentation and classification. The outcome was computed over the recombination of the results from all 16 sub-images for each raw image. The results are shown in Fig. 6 and Table 2. The mean square error obtained with deep learning-based classification was as low as 0.0622 using WS-CNN, 0.0677 using UN-CNN, and 0.0764 using UW-CNN. The CNN-based

Table 2. RSEs of the measured viability using different segmentation/classification methods.

Methods	Mean	Std Dev
UW-CNN	0.076	0.090
UW-SVM	0.104	0.129
UW-KNN	0.151	0.127
UN-CNN	0.067	0.071
UN-SVM	0.096	0.111
UN-KNN	0.152	0.131
WS-CNN	0.062	0.060
WS-SVM	0.111	0.119
WS-KNN	0.129	0.115

classification demonstrated the best accuracy for viability analysis, regardless of the segmentation algorithms. Additionally, deep learning CNN classifier-based viability analysis was better than those feature-based machine learning classifiers such as SVM and KNN. Figure 7 shows typical image outputs of the segmentation and classification processes using UN-CNN, UW-CNN and WS-CNN. In the output images, segmented chondrocyte areas were circled with different colors: red for dead cells, green for live cells, yellow for cell clusters, and magenta for no-cell areas.

4. Discussion

Inspired by the deep watershed transform [18], in this report, we performed a second segmentation with the watershed method on the output of the deep learning segmentation. The watershed targeted the margin features and significantly improved the touching problem. The F1 score reached over 0.9 on the chondrocyte segmentation in articular cartilage. This strategy may be quite useful in other biological and clinical image processing methods when higher accuracy is sought in segmentation. For example, this strategy can improve the segmentation when cells are densely packed in cell culture.

In our cell viability analysis, TPAF/SHG images are first converted to RGB color images. The visual assessment of these images provides a reliable way to separate dead/live populations. However, conventional machine learning methods such as SVM and KNN did not perform well in the cell classification, partly because these methods used pre-defined features that might not completely cover the features that the human brain uses to process the images. Deep learning networks may establish much more complicated models through the training process although the exact model is inscrutable. In our study, we demonstrated that CNN outperformed SVM and KNN methods in our tasks. In our classification, we used two 5-category multivariate classifications, one for live cell ($n=0-4$) and the other for total cell ($n=0-4$) counts. As shown in Fig. 2(B), it is possible to perform one 15-category classification; however, too many categories in classification reduces the accuracy and increases the training time significantly. We found that the two 5-category classifications were more efficient than the single 15-category classification. This suggests, as shown in the results, that the dual classifications are the choice to improve the accuracy of the viability analysis while maintaining a low computing cost.

In deep learning neural networks, we used U-Net and CNN (AlexNet) neural networks, but there were other similar networks, such as R-CNN [28], Inception [29], ResNet [30], etc., which use many more layers for image segmentation/detection and classification. Deeper networks may increase the performance and quantitative accuracy but with increased computing costs and high reliance on the computing capability of processors (CPUs, single GPU, or multiple GPUs). For instance, on the same classification training dataset, batch size, and epoch length, the training time of ResNet with hundreds of neural layers was 2-3 times longer than AlexNet [31]. In this paper, the training time of developed models (U-Net and AlexNet architectures) cost only less than half an hour for thousands of images. Nevertheless, in future work, deeper networks will be explored with a balance of the depth and the computing cost to improve the accuracy further.

Additionally, nonlinear optical microscopy (NLOM) imaged cell function and morphology and indicated the health status of cartilage through 3D z-stack TPAF/SHG images. However, we built deep learning models in which 2D TPAF/SHG images and detection/classification annotations were used as training inputs. Lack of the correlation between slices may induce low sensitivity and specificity because parts of cellular volume miss morphological and functional features associated with the whole volume. Moreover, it has been demonstrated that 3D deep learning models may increase performance and accuracy by providing more z-stack information. For instance, Ronneberger et al. demonstrated the performance of confocal kidney segmentation of the 3D architecture outperforms equivalent 2D implementation [32]. Meanwhile, in the classification of stack images, 3D models such as multi-view CNN and volumetric CNN were developed. For instance, Guibas et al. demonstrated that volumetric CNN is significantly worse than multi-view

CNN [33] even though their inputs have similar amounts of information. Based on this problem, the group proposed two new architectures of volumetric CNNs, which outperform state-of-the-art volumetric CNNs, achieving comparable performance to multi-view CNNs at the same resolution. It was demonstrated in the classification of phase-contrast image sequences of stem cells that volumetric CNN can outperform the 2D CNN features and the popular hand-crafted features for visual representation [34]. But, in PET images of human brains, the classification of Alzheimer's disease by using 3D-CNN and 2D-CNN is not significantly different [35]. Thus, the classification performance of 3D-CNN and 2D-CNN is highly dependent on the types of neural networks and image features. In the future, we will compare the performance of 2D and 3D neural networks with respect to the chondrocyte viability assessment.

5. Conclusion

In this paper, the developed multivariate deep learning algorithm showed the capability of analyzing cell viability from nonlabelled cellular images acquired from a nonlinear optical microscope. A new strategy counted all touching cells with high accuracy by taking advantage of complicated multivariate classification. Chondrocytes were detected and classified with 90% accuracy using deep learning methods. The segmentation and classification strategy described in this report provided a high accuracy, automated method for the chondrocyte viability analysis. We anticipate that this automated imaging processing method will have an impact on *in vivo* or *ex vivo* studies of cartilage degeneration. Additionally, the strategy presented here may provide a concept to overcome the cell touching patterns for cell counting in similar situations.

Funding. Chan Zuckerberg Initiative (2019-198009); National Institute of General Medical Sciences (P20 GM121342); National Institute of General Medical Sciences (P20 GM103499); MTF Biologics (Extramural Research Grant); National Science Foundation (1539034); National Cancer Institute (P30 CA138313).

Acknowledgments. This research was supported by South Carolina IDeA Networks of Biomedical Research Excellence (SC INBRE), a National Institutes of Health (NIH) funded center (Award P20 GM103499), MTF Biologics Extramural Research Grant, South Carolina Translation Research Improving Musculoskeletal Health (TRIMH), an NIH funded Center of Biomedical Research Excellence (Award P20GM121342), Clemson University's Robert H. Brooks Sports Science Institute (RHBSSI) Seed Grant, a grant from the National Science Foundation (1539034), and a grant from Chan Zuckerberg Initiative DAF to T.P. (2019-198009). The imaging experiments were performed on the commercial microscopes in the user facility supported by Cell & Molecular Imaging Shared Resource, Hollings Cancer Center, Medical University of South Carolina (P30 CA138313). This work was supported in part by the National Science Foundation EPSCoR Program under NSF Award #OIA-1655740.

Disclosures. The authors declare that there are no conflicts of interest related to this article.

Supplemental document. See [Supplement 1](#) for supporting content.

References

1. P. G. Bush, P. D. Hodkinson, G. L. Hamilton, and A. C. Hall, "Viability and volume of in situ bovine articular chondrocytes—changes following a single impact and effects of medium osmolarity," *Osteoarthritis and Cartilage* **13**(1), 54–65 (2005).
2. L. V. Gulotta, J. R. Rudzki, D. Kovacevic, C. C. T. Chen, D. Milentijevic, and R. J. Williams, "Chondrocyte death and cartilage degradation after autologous osteochondral transplantation surgery in a rabbit model," *American J. Sports Med.* **37**, 1324–1333 (2009).
3. Y. Li, X. Chen, B. Watkins, N. Saini, S. Gannon, E. Nadeau, R. Reeves, B. Gao, V. Pelligrini, H. Yao, J. Mercuri, and T. Ye, "Nonlabeling and quantitative assessment of chondrocyte viability in articular cartilage with intrinsic nonlinear optical signatures," *Exp. Biol. Med.* **245**(4), 348–359 (2020).
4. T. Liu, E. Jurrus, M. Seyedhosseini, M. Ellisman, and T. Tasdizen, "Watershed merge tree classification for electron microscopy image segmentation," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)* (2012), pp. 133–137.
5. T. Kanade, Z. Yin, R. Bise, S. Huh, S. Eom, M. F. Sandbothe, and M. Chen, "Cell image analysis: algorithms, system and applications," in *2011 IEEE Workshop on Applications of Computer Vision (WACV)* (2011), pp. 374–381.
6. B. J. Erickson, P. Korfiatis, Z. Akkus, and T. L. Kline, "Machine learning for medical imaging," *RadioGraphics* **37**(2), 505–515 (2017).
7. S. K. Zhou, H. Greenspan, and D. Shen, *Deep Learning for Medical Image Analysis* (Academic Press, 2017).

8. T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, A. Dovzhenko, O. Tietz, C. Dal Bosco, S. Walsh, D. Saltukoglu, T. L. Tay, M. Prinz, K. Palme, M. Simons, I. Diester, T. Brox, and O. Ronneberger, "U-Net: deep learning for cell counting, detection, and morphometry," *Nat. Methods* **16**(1), 67–70 (2019).
9. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds., Lecture Notes in Computer Science (Springer International Publishing, 2015), pp. 234–241.
10. T. C. Hollon, B. Pandian, A. R. Adapa, E. Urias, A. V. Save, S. S. S. Khalsa, D. G. Eichberg, R. S. D'Amico, Z. U. Farooq, S. Lewis, P. D. Petridis, T. Marie, A. H. Shah, H. J. L. Garton, C. O. Maher, J. A. Heth, E. L. McKean, S. E. Sullivan, S. L. Hervey-Jumper, P. G. Patil, B. G. Thompson, O. Sagher, G. M. McKhann, R. J. Komotar, M. E. Ivan, M. Snuderl, M. L. Otten, T. D. Johnson, M. B. Sisti, J. N. Bruce, K. M. Muraszko, J. Trautman, C. W. Freudiger, P. Canoll, H. Lee, S. Camelo-Piragua, and D. A. Orringer, "Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks," *Nat. Med.* **26**(1), 52–58 (2020).
11. M. Kowal, M. Žejmo, M. Skobel, J. Korbicz, and R. Monczak, "Cell nuclei segmentation in cytological images using convolutional neural network and seeded watershed algorithm," *J Digit Imaging* **33**(1), 231–242 (2020).
12. Y. Yu, J. Wang, C. W. Ng, Y. Ma, S. Mo, E. L. S. Fong, J. Xing, Z. Song, Y. Xie, K. Si, A. Wee, R. E. Welsch, P. T. C. So, and H. Yu, "Deep learning enables automated scoring of liver fibrosis stages," *Sci. Rep.* **8**(1), 16016 (2018).
13. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: an open-source platform for biological-image analysis," *Nat. Methods* **9**(7), 676–682 (2012).
14. M. Abdolhoseini, M. G. Kluge, F. R. Walker, and S. J. Johnson, "Segmentation of heavily clustered nuclei from histopathological images," *Sci. Rep.* **9**(1), 4551 (2019).
15. C. Ounkomol, S. Seshamani, M. M. Maleckar, F. Collman, and G. R. Johnson, "Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy," *Nat. Methods* **15**(11), 917–920 (2018).
16. E. Moen, D. Bannon, T. Kudo, W. Graf, M. Covert, and D. Van Valen, "Deep learning for cellular image analysis," *Nat. Methods* **16**(12), 1233–1246 (2019).
17. M. Gupta, *Cell Identification by Blob Detection* (2012), p. 236.
18. M. Bai and R. Urtasun, "Deep watershed transform for instance segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2017), pp. 5221–5229.
19. X. Chen, Y. Li, N. Wyman, Z. Zhang, H. Fan, M. Le, S. Gannon, C. Rose, Z. Zhang, J. Mercuri, H. Yao, B. Gao, S. Woolf, T. Pecot, and T. Ye, "Cartilage-Net," Github 2021 <https://github.com/chenxun511happy/Cartilage-Net>.
20. M. P. Sampat, A. C. Bovik, J. K. Aggarwal, and K. R. Castleman, "Supervised parametric and non-parametric classification of chromosome images," *Pattern Recognition* **38**(8), 1209–1223 (2005).
21. M. Wang, X. Zhou, F. Li, J. Huckins, R. W. King, and S. T. C. Wong, "Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy," *Bioinformatics* **24**(1), 94–101 (2008).
22. N. Coudray, P. S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, A. L. Moreira, N. Razavian, and A. Tsirigos, "Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning," *Nat. Med.* **24**(10), 1559–1567 (2018).
23. A. B. L. Larsen, A. B. Dahl, and R. Larsen, "Oriented shape index histograms for cell classification," in *Image Analysis*, R. R. Paulsen and K. S. Pedersen, eds., Lecture Notes in Computer Science (Springer International Publishing, 2015), pp. 16–25.
24. A. Gençtav, S. Aksoy, and S. Önder, "Unsupervised segmentation and classification of cervical cell images," *Pattern Recognition* **45**(12), 4151–4168 (2012).
25. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, and D. Cournapeau, "Scikit-learn: Machine Learning in Python," *MACHINE LEARNING IN PYTHON 6* (n.d.).
26. J. Li and J. P. Fine, "ROC analysis with multiple classes and multiple tests: methodology and its application in microarray studies," *Biostatistics* **9**(3), 566–576 (2008).
27. "Receiver Operating Characteristic (ROC) — scikit-learn 0.24.1 documentation," https://scikit-learn.org/stable/auto_examples/model_selection/plot_roc.html.
28. A. O. Vuola, S. U. Akram, and J. Kannala, "Mask-RCNN and U-Net ensemble for nuclei segmentation," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (2019), pp. 208–212.
29. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015), pp. 1–9.
30. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2016), pp. 770–778.
31. Y. You, Z. Zhang, C.-J. Hsieh, and J. Demmel, "100-epoch ImageNet Training with AlexNet in 24 Minutes," ArXiv (2017).
32. Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, eds., Lecture Notes in Computer Science (Springer International Publishing, 2016), pp. 424–432.

33. C. R. Qi, H. Su, M. Niessner, A. Dai, M. Yan, and L. J. Guibas, "Volumetric and multi-view CNNs for object classification on 3D data," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2016), pp. 5648–5656.
34. W.-Z. Nie, W.-H. Li, A.-A. Liu, T. Hao, and Y.-T. Su, "3D convolutional networks-based mitotic event detection in time-lapse phase contrast microscopy image sequences of stem cell populations," in *Proceeding of the 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, 2016), pp. 55–62.
35. D. Cheng and M. Liu, "Classification of Alzheimer's disease by cascaded convolutional neural networks using PET images," in *Machine Learning in Medical Imaging*, Q. Wang, Y. Shi, H.-I. Suk, and K. Suzuki, eds., Lecture Notes in Computer Science (Springer International Publishing, 2017), pp. 106–113.